

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 04-125599
 (43)Date of publication of application : 27.04.1992

(51)Int.CI. G10L 3/00

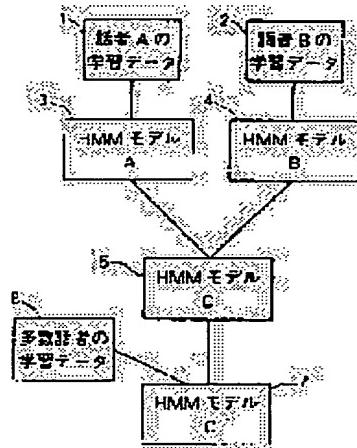
(21)Application number : 02-246863 (71)Applicant : NEC CORP
 (22)Date of filing : 17.09.1990 (72)Inventor : ISOTANI RYOSUKE

(54) REFERENCE PATTERN GENERATING METHOD

(57)Abstract:

PURPOSE: To easily decide the parameter of a reference pattern by finding vector output probability distribution represented in mixed continuous distribution by synthesizing the vector output probability distribution of plural learned reference patterns.

CONSTITUTION: An HMM(imbedded Markov) model A(3) is generated from the learning data (1) of a talker A, and an HMM model B(4) from the learning data (2) of a talker B. An HMM model C(5) for speech recognition of unspecific talker is generated from the models A and B. The model (C) can be used as the HMM model for recognition of unspecific talker as it is, and also, a better model C'(7) can be generated by using the learning data (6) for large number of talkers. In such a way, it is possible to easily decide the parameter of the reference pattern in which the vector output probability is represented in the mixed continuous distribution by using the plural learned reference patterns.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C) 1998,2003 Japan Patent Office

⑩ 日本国特許庁 (JP) ⑪ 特許出願公開
⑫ 公開特許公報 (A) 平4-125599

⑬ Int. Cl.⁵
G 10 L 3/00

識別記号 301 E
301 C

厅内整理番号 8842-5H
8842-5H

⑭ 公開 平成4年(1992)4月27日

審査請求 未請求 請求項の数 3 (全4頁)

⑮ 発明の名称 標準パターン作成方法

⑯ 特 願 平2-246863
⑰ 出 願 平2(1990)9月17日

⑱ 発明者 碓谷 充輔 東京都港区芝5丁目7番1号 日本電気株式会社内
⑲ 出願人 日本電気株式会社 東京都港区芝5丁目7番1号
⑳ 代理人 弁理士 岩佐 義幸

明細書

1. 発明の名称

標準パターン作成方法

2. 特許請求の範囲

(1) 状態の集合と状態間の遷移確率と状態あるいは遷移のベクトル出力確率とによって定義される標準パターンの作成方法において、

ベクトル出力確率が連続分布で表される複数の標準パターンの対応する状態あるいは遷移のベクトル出力確率分布を重み付きで混合した混合連続分布を状態あるいは遷移のベクトル出力確率とする標準パターンを作成することを特徴とする標準パターン作成方法。

(2) 状態の集合と状態間の遷移確率と状態あるいは遷移のベクトル出力確率とによって定義される音声認識用の標準パターンの作成方法において、複数の話者について話者ごとにその話者の音声データを用いて学習して作成されたベクトル出力確率が連続分布で表される標準パターンの対応する状態あるいは遷移のベクトル出力確率分布を重

み付きで混合した混合連続分布を状態あるいは遷移のベクトル出力確率とする標準パターンを作成することを特徴とする標準パターン作成方法。

(3) 状態の集合と状態間の遷移確率と状態あるいは遷移のベクトル出力確率とによって定義される音声認識用の標準パターンの作成方法において、異なる環境で発声あるいは収録した音声データを用いて環境ごとに学習して作成されたベクトル出力確率が連続分布で表される標準パターンの対応する状態あるいは遷移のベクトル出力確率分布を重み付きで混合した混合連続分布を状態あるいは遷移のベクトル出力確率とする標準パターンを作成することを特徴とする標準パターン作成方法。

3. 発明の詳細な説明

【産業上の利用分野】

本発明は、音声認識等パターン認識に用いられる標準パターンの作成方法に関する。

【従来の技術】

音声認識などパターン認識の分野で、認識用の標準パターンとして確率モデルを用いる方法が近

年注目されており、特に隠れマルコフモデル（以下HMMと呼ぶ）は音声認識の分野で標準パターンを表すモデルとして広く用いられている。

HMMは状態の集合と状態間の遷移確率と状態あるいは遷移のベクトル出力確率によって定義され、入力パターンに対する各HMMの尤度を計算することにより認識を行う。HMMによる音声認識については、刊行物「確率モデルによる音声認識」中川聖一著に詳しく述べられている。

各状態（あるいは遷移）のベクトル出力確率が混合連続分布で表されるHMMモデルのパラメータを決定する方法として、Baum-Welchアルゴリズムなど、ある初期値から学習用データを用いてパラメータを繰り返し更新する学習法が知られている。この場合、出力確率分布の平均値などのパラメータの初期値は、混合する各分布毎に決定する必要がある。これらのパラメータの初期値を与える方法としては、

(a)乱数で与える

(b)単一の分布の場合のパラメータに乱数値でほか

を用いる場合、効率よく学習が行われるために初期値の設定が重要であることが知られているが、(a)のように乱数を用いたり(b)のように単一の分布の場合のパラメータを用いるのでは、学習の収束までに時間がかかり、また収束値も全体の最適値ではなく局所的な最適値になる可能性が高い。一方(c)の方法は、パラメータ更新のための学習を必ずしも必要とせず、また、更新の初期値として用いる場合でも少ない繰り返し回数で収束すると考えられるが、クラスタリングのための計算などが必要で、計算量が多くなるという欠点があった。

本発明の目的は、このような欠点を解消した標準パターン作成方法を提供することにある。

〔課題を解決するための手段〕

第1の発明は、状態の集合と状態間の遷移確率と状態あるいは遷移のベクトル出力確率とによって定義される標準パターンの作成方法において、

ベクトル出力確率が連続分布で表される複数の標準パターンの対応する状態あるいは遷移のベクトル出力確率分布を重み付きで混合した混合連続

し作用を行う（「連続出力分布型HMMによる日本語音韻認識の検討」）電子情報通信学会音声研究会資料SP89-48）

などの方法が知られている。

一方、ある初期値から更新によって求めるではなく学習データから直接パラメータを決定する方法として、

(c)学習データをセグメンテーションしたあとクラスタリングを行って混合する分布数のクラスターを求め、各クラスターのデータから平均値等のパラメータを求める方法が知られている（“High Performance Connected Digit Recognition Using Hidden Markov Models”, IEEE Transaction on Acoustics, Speech, and Signal Processing, Vol.37, No.8, pp.1214-1224, August 1989）。

このようにして決められた値を初期値として、 Baum-Welchアルゴリズムなどにより更新を行うことができる。

〔発明が解決しようとする課題〕

学習により繰り返しパラメータを更新する方法

分布を状態あるいは遷移のベクトル出力確率とする標準パターンを作成することを特徴とする。

第2の発明は、状態の集合と状態間の遷移確率と状態あるいは遷移のベクトル出力確率とによって定義される音声認識用の標準パターンの作成方法において、

複数の話者について話者ごとにその話者の音声データを用いて学習して作成されたベクトル出力確率が連続分布で表される標準パターンの対応する状態あるいは遷移のベクトル出力確率分布を重み付きで混合した混合連続分布を状態あるいは遷移のベクトル出力確率とする標準パターンを作成することを特徴とする。

第3の発明は、状態の集合と状態間の遷移確率と状態あるいは遷移のベクトル出力確率とによって定義される音声認識用の標準パターンの作成方法において、

異なる環境で発声あるいは収録した音声データを用いて環境ごとに学習して作成されたベクトル出力確率が連続分布で表される標準パターンの対

特開平4-125599(3)

応する状態あるいは遷移のベクトル出力確率分布を重み付きで混合した混合連続分布を状態あるいは遷移のベクトル出力確率とする標準パターンを作成することを特徴とする。

(作用)

本発明によれば、混合連続分布で表されるベクトル出力確率分布を、すでに学習済みの複数の標準パターンのベクトル出力確率分布から合成して求めることにより、標準パターンのパラメータを簡易に決定することができる。また、合成に用いる標準パターンを適切に選べば、Baum-Welch法などの学習の初期パラメータとして用いる場合、乱数で初期パラメータを決定する場合などに比べ少ない学習回数で収束し、局所的な最適値に収束する確率も小さくなると期待される。また、学習によるパラメータ更新を行わずそのまま用いることもできる。

第2の発明のように、合成に複数の話者について話者ごとにその話者の音声データを用いて学習して作成された標準パターンを用いれば、ベクト

ル出力確率が混合連続出力分布で表される不特定話者音声認識用の標準パターンを簡易に作成することができる。

第3の発明のように、合成に異なる環境で発声あるいは収録した音声データを用いて環境ごとに学習して作成された標準パターンを用いれば、ベクトル出力確率が混合連続出力分布で表される環境の変動に強い標準パターンを簡易に作成することができる。

(実施例)

第1図は、第1の発明を不特定話者音声認識用のHMMモデル作成に適用した実施例を説明するためのブロック図である。話者Aの学習データ(1)からHMMモデルA(3)を、話者Bの学習データ(2)からHMMモデルB(4)を作成する。話者A、Bとしては、たとえば男性、女性から標準的な話者を1名ずつ選んで用いる。HMMモデルは第2図に示すような形のモデルとする。各状態*i*に対し、状態遷移確率 $a_{ii}^1, a_{ii+1}^1 (a_{ii}^1 + a_{ii+1}^1 = 1)$ と出力ベクトル y に対する出力確率分布 $b_i^1(y)$

が定められている。モデルAの状態遷移確率、出力確率分布を、それぞれ $a^{1*}, b^{1*}(y)$ などと表す。出力ベクトル確率分布が単一ガウス分布で表されたとすると、

$$b_i^{1*}(y) = N(y, \mu_i^{1*}, \Sigma_i^{1*})$$

$$b_i^{2*}(y) = N(y, \mu_i^{2*}, \Sigma_i^{2*})$$

と表される。ここで、 $N(y, \mu_i, \Sigma_i)$ は平均ベクトルを μ_i 、共分散行列を Σ_i とする多次元ガウス分布を表す。モデルAとモデルBから、不特定話者音声認識用のHMMモデルC(5)を作成する。モデルCの状態遷移確率を a_{ii}^c, a_{ii+1}^c 、出力確率分布を b_i^c とする。出力確率分布が、次のような混合数2の混合ガウス分布で表されるとする。

$$b_i^c(y) = \lambda^1 N(y, \mu_i^1, \Sigma_i^1) + \lambda^2 N(y, \mu_i^2, \Sigma_i^2)$$

このとき、モデルCの各パラメータを次のように定める。

$$a_{ii}^c = (a_{ii}^{1*} + a_{ii}^{2*}) / 2$$

$$a_{ii+1}^c = (a_{ii+1}^{1*} + a_{ii+1}^{2*}) / 2$$

$$\mu_i^1 = \mu_i^{1*}, \Sigma_i^1 = \Sigma_i^{1*}$$

$$\mu_i^2 = \mu_i^{2*}, \Sigma_i^2 = \Sigma_i^{2*}$$

$$\lambda^1 = \lambda^2 = 1/2$$

このようにして作成されたモデルCは、そのまま不特定話者音声認識用のHMMモデルとして用いることもでき、また、さらに多数の話者の学習データ(6)を用いてBaum-Welch法などで学習を行い、よりよいモデルC'(7)を作成するための初期モデルとして用いることもできる。

モデルA、Bとして出力確率分布が混合ガウス分布で表されるものが用意されている場合にも、同様にモデルCを作成することができる。この場合、モデルCの出力確率分布の混合数は、モデルA、Bの出力確率分布の混合数の和になる。

次に、第2の発明の一実施例について説明する。多数の話者が発声した少數語彙の音声データをクラスタリングすることにより話者をM個のクラスタに分け、各クラスタからクラスタ中心の話者M名を選ぶ。M名の各話者について、HMM学習に必要な量の音声データをもとに、出力確率分布が

特開平4-125599(4)

単一ガウス分布で表されるHMMモデルを学習して作成する。作成されたM個のモデルから、第1の発明の実施例と同様に混合数がMの混合ガウス分布を出力確率分布とするHMMモデルを作成することにより不特定話者音声認識用のHMMモデルが得られる。M名の話者を選ぶためのクラスタリングに用いるデータは少數のデータでよいので、従来の技術の(c)に比べ計算量は少なくなる。

最後に、第3の発明の一実施例について説明する。第1の発明の実施例において、モデルA、Bの選び方として、ある話者の異なる環境下（たとえば、静かな環境と雑音の多い環境）で発声したデータを用いて学習したモデルを用いれば、モデルCとして環境の変動に強い認識モデルを作成することができる。

〔発明の効果〕

以上述べたように、第1の発明によれば、すでに学習されている複数の標準パターンを用いて、ベクトル出力確率が混合連続分布で表される標準パターンのパラメータを簡単に決定することができ

き、そのまま、あるいはこの値を初期値とした少數回の学習でパターン認識に用いることができる。また、第2、第3の発明によれば、不特定話者用、環境の変動に強い標準パターンをそれぞれ簡易に作成することができる。

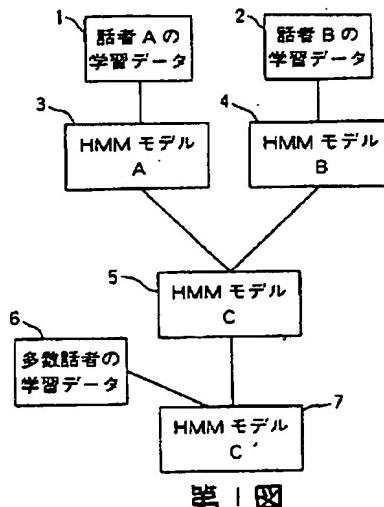
4. 図面の簡単な説明

第1図は、第1の発明を不特定話者音声認識用のHMMモデル作成に適用した実施例を説明するためのブロック図。

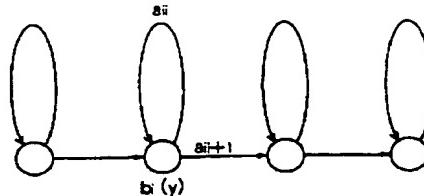
第2図は、実施例におけるHMMモデルの形を示す図である。

- 1 ····· 話者Aの学習データ
- 2 ····· 話者Bの学習データ
- 3 ····· HMMモデルA
- 4 ····· HMMモデルB
- 5 ····· HMMモデルC
- 6 ····· 多数話者の学習データ
- 7 ····· HMMモデルC'

代理人 弁理士 岩佐義幸



第1図



第2図